

A Correction Note for Price Dynamics in a Markovian Limit Order Market

Geliang Zhang^{†§}, Hugh Christensen[†], Guolong Li[‡], and Simon Godsill[†]

Abstract. A previously published and widely quoted paper, Price Dynamics in a Markovian Limit Order Market, has provided analytic solutions for many interesting quantities for the price dynamics of the limit order book. However, for unbalanced order flow cases, one of its major results, the price increase probability conditional on the state of the limit order book is found to be incorrect due to its citing an erroneous formula from another paper. We correct this error by proposing another analytical solution using the characteristic functions of probability distributions. The final analytical solution presented in this paper is of a simple form and easy to calculate.

Key words. limit order book, random walk, price dynamics, hitting probability

AMS subject classifications. 60J27, 60E10, 60K25

1. Introduction. The limit order book (LOB) is a depository of resting limit orders that provides liquidity allowing participants to trade in the market. Much effort has been put into trying to extract information from the LOB and predict the direction of price change in the short future.

A recent paper [2] showed that several interesting conditional probabilities, including the direction of the next price change conditional on the state of the LOB, can be obtained analytically using a simplified model for the LOB based on a previous model proposed in [4]. This simplified model can capture the dynamics of market orders and limit orders and their influence on price dynamics in a more analytically tractable way. The paper brought new insights into the LOB, helped people understand its dynamics [1],[8],[6], and inspired applications based on it, for example, building optimal trading strategies [9].

However, we find that one of its major results, Proposition 3, which provides the analytic solution to the conditional probability of price movement of the LOB when the order flow is unbalanced, is incorrect. This proposition itself is important because nowadays much attention has been focused on the study of the order flow imbalance (OFI) process, which is thought to be more sensitive to market information than the price [3][10]. And Proposition 3 gives the short term price dynamics of the LOB driven by this OFI process under the Markovian assumption in [2].

A numerical and mathematical proof for this mistake is provided in this correction note, as well as a corrected analytic formula for the conditional probability of the upward price movement.

1.1. Summary. The model presented in [2], denoted as Model 2013, has attracted broad attention since its publication due to its ability to capture many characteristics of the limit order book while being analytic tractability. In this correction note, we propose a correct

[†]Signal Processing and Communication Laboratory, Engineering Department, University of Cambridge, CB2 1PZ, UK. (gz246@cam.ac.uk, hlc54@cam.ac.uk, sjg30@cam.ac.uk)

[‡]Faculty of Mathematics, University of Cambridge, CB3 0WA, UK. (gl288@cam.ac.uk)

[§]Thanks to Cambridge Commonwealth, European and International Trust to provide financial support.

analytical solution to replace the original Proposition 3 in [2] for Model 2013, using a first passage time approach. The correction takes the form of the inverse of a characteristic function, which is simple and can be evaluated fast and easily. The probability values obtained using our solution are found to be consistent with values obtained using the Laplace transform (LT) methods used in another previous paper [4]. The first passage time argument itself is standard and can be found in other literature [7]. The analytical solution presented here has the advantage of simplicity, while the LT methods involve the usage of numerical inverse Laplace transforms (ILT) and infinite continued fractions, which add significant complexity in computation for real-world applications.

1.2. Outline. The paper is organized as follows. Section 2 illustrates where the error occurs. Section 3 presents our analytic solution for the probability of upward movement in the price in Model 2013 to replace the incorrect Proposition 3 in [2].

2. Problem.

2.1. Probability of mid-price increase. The conditional probability of a mid-price increase can be reformulated into the probability of hitting the boundary in a 2-dimension random walk problem (see section 3.3 of [2]). To give a more detailed description, mid-price variations are driven by the four independent order flows listed below:

1. Limit orders at the ask arrive as a Poisson process with intensity λ^a .
2. Market orders and cancellations at the ask arrive as a Poisson process with intensity $\mu^a + \theta^a$.
3. Limit orders at the bid arrive as a Poisson process with intensity λ^b .
4. Market orders and cancellations at the bid arrive as a Poisson process with intensity $\mu^b + \theta^b$.

Limit orders increase the number of orders in the LOB, while market orders and cancellation orders decrease the number of orders. If orders on either side (bid or ask) are depleted, then the mid-price will change. With another parameter $\Lambda = \lambda^a + \mu^a + \theta^a + \lambda^b + \mu^b + \theta^b$, the dynamics of bid and ask queues, denoted as q_t , may be then represented as follows:

$$(2.1) \quad q_t = M_{N_{\Lambda t}}, \text{ for } \Lambda = \lambda^a + \mu^a + \theta^a + \lambda^b + \mu^b + \theta^b.$$

$N_{\Lambda t}$ is a Poisson process with intensity Λ . $(M_n, n \geq 0)$ is a random walk on \mathbb{N}^2 killed when it hits the boundary, where the transition probabilities are

$$(2.2) \quad p_{0,1} = \frac{\lambda^a}{\Lambda}, \quad p_{1,0} = \frac{\lambda^b}{\Lambda}, \quad p_{0,-1} = \frac{\mu^a + \theta^a}{\Lambda}, \quad p_{-1,0} = \frac{\mu^b + \theta^b}{\Lambda}.$$

If at a given time t , the limit order book has n orders at the bid side, and p orders at the ask side, then Proposition 3 of [2] states that the probability $p^{up}(n, p)$ that, the next price

move is an increase can be obtained using the following equations:

$$\begin{aligned}
 (2.3) \quad p^{up}(n, p) = & 1 - \frac{1}{\pi} \left(\frac{\mu^a + \theta^a}{\lambda^a} \right)^p \frac{2[\lambda^a (\mu^a + \theta^a)]^{1/2}}{\mu^a + \theta^a + \lambda^a} \\
 & \times \int_0^\pi \frac{2\lambda^b Z(t) - (\Lambda - 2[\lambda^a (\mu^a + \theta^a)]^{1/2} \cos(t))}{\frac{2[(\mu^a + \theta^a)\lambda^a]^{1/2}}{\lambda^b + \mu^b + \theta^b} \cos(t) - 1} \\
 & \times \frac{Z(t)^n \sin(pt) \sin(t)}{\sqrt{(\Lambda - 2[(\mu^a + \theta^a)\lambda^a]^{1/2} \cos(t))^2 - 4(\mu^b + \theta^b)\lambda^b}} dt,
 \end{aligned}$$

where $(Z(t), t \geq 0)$ is defined for $t > 0$ by

$$\begin{aligned}
 (2.4) \quad Z(t) = & \{\Lambda - 2[(\mu^a + \theta^a)\lambda^a]^{1/2} \cos(t) \\
 & - \sqrt{(\Lambda - 2[(\mu^a + \theta^a)\lambda^a]^{1/2} \cos(t))^2 - 4(\mu^b + \theta^b)\lambda^b}\} / 2\lambda^a.
 \end{aligned}$$

2.2. Comments on Model 2013. A final analytical solution for the probability of a price increase becomes possible using results related to the 2-D random walk theory. The final solution with such forms in (2.3) and (2.4) is easy to evaluate.

However, we find there are serious errors in its conclusion. The results obtained in [2] make us unable to use this random walk approach to obtain the correct probability of mid-price increase in the analytic form given. The errors were firstly observed by the authors when empirical simulations kept giving negative probabilities, indicating that analytic results in these equations have not been derived correctly. After further research we found that the non-convergent term, $\left(\frac{\mu^a + \theta^a}{\lambda^a}\right)^p$ in (2.3), comes from another incorrectly-derived theory published in another paper, thus making it difficult to correct by just simply replace it with a correct equation. The incorrectly-derived theory which [2] cited tried to obtain its conclusion by generalizing an already-solved Riemann-boundary value problem for non-negative mean drift cases of the 2-D random walk (see [11]) to negative-drift 2-D random walk cases (see the Proof of Theorem 6 of [13]). It seems that such a generalization can not be obtained as claimed, since it produces negative values for probability functions. Thus we have to work it out using an alternative way, via characteristic function of probabilities.

Some parameter errors possibly caused by typos also exist in Proposition 3 of [2], shown here as in (2.3) and (2.4). But they are of little importance compared with the fundamental error.

2.3. Discussion of mathematical reasons for the error. The drifts for two random walks in the x and y-axis are defined as $\sum_{-1 \leq i, j \leq 1} i p_{i,j}$ and $\sum_{-1 \leq i, j \leq 1} j p_{i,j}$ respectively. For positive-drift and zero-drift cases, we have $\frac{p_{0,-1}}{p_{0,1}} \leq 1$. However, for negative-drift cases, we have $\frac{p_{0,-1}}{p_{0,1}} > 1$ since $\sum_{-1 \leq i, j \leq 1} j p_{i,j} = -p(0, -1) + p(0, 1) < 0$ using the fact that $p_{i,j} = 0$ except for $p_{0,1}, p_{0,-1}, p_{-1,0}$ and $p_{1,0}$. Therefore, for negative-drift cases, Theorem 6 of [13] can not stand because of the non-convergent term of $\left(\frac{p_{0,-1}}{p_{0,1}}\right)^{j_0/2}$ in the hitting probabilities $h^{i_0, j_0}(x=1)$ generated by $h^{i_0, j_0}(x)$. The derivation of equation (3.10) of [13] is based on

the strict condition that it is an non-negative drift random walk case, and can not be easily generalized by just choosing a different w function.

And for the case in which the authors in [2] try to apply this result, we can see that $\frac{p_{0,-1}}{p_{0,1}} = \frac{\mu^a + \theta^a}{\lambda^a} > 1$, which is a negative-drift case. Therefore, the authors in [2] cited a result which does not hold, as was claimed, in negative-drift cases, and thus they obtained a wrong analytic result.

3. Corrected analytical solution using characteristic functions. In this section, we propose another approach using characteristic functions of the probability of interest to obtain the probability of mid-price increase. The approach taken in this section is standard can be found in literature such as [7].

3.1. Theoretic solution. Using the same framework as proposed in [2], we consider the price dynamics as a simple random walk in the quarter plane with four probabilities $p_{0,1}$, $p_{0,-1}$, $p_{-1,0}$ and $p_{1,0}$ defined as in (2.2). What we are interested in is the probability of point starting at (x_0, y_0) hitting the X-axis earlier than the Y-axis following the simple random walk mentioned above. Then the process of the point's x-value and the process of its y-value will be independent. It hits the X-axis first if its y-value goes to 0 first. So, let $T_X(x_0)$ be the time its x-value starting at x_0 goes to 0 and let $T_Y(y_0)$ be the time its y-value starting at y_0 goes to 0. Therefore the problems will be solved if we know the probability distribution of $T_X(x_0)$ and $T_Y(y_0)$; in this case we can then obtain the probability distribution of $T_Y(y_0) < T_X(x_0)$.

In order to work out the distribution of $T_Y(y_0)$, we look at its characteristic function, i.e. $\Phi(u) = E(\exp^{iuT_Y(y_0)})$ for some real u . We can get the probability density function from this characteristic function by Fourier Transform. Let $t_Y(1)$ be the time the y-value hits 1, $t_Y(2)$ be the time the y-value hits 2 and so on. Note that $t_Y(2) - t_Y(1)$, $t_Y(3) - t_Y(2)$, ... are all independent and identically distributed with each having the same distribution as $T_Y(1)$. Moreover $T_Y(y_0) = t_Y(y_0) - t_Y(y_0 - 1) + t_Y(y_0 - 1) - t_Y(y_0 - 2) \dots + t_Y(1) - t_Y(0)$, and by above definitions we have $T_Y(1) = t_Y(1) - t_Y(0)$. So,

$$\begin{aligned}
 E(\exp^{iuT_Y(y_0)}) &= E(\exp^{iu[t_Y(y_0) - t_Y(y_0 - 1) + t_Y(y_0 - 1) - t_Y(y_0 - 2) \dots + t_Y(1) - t_Y(0)]}) \\
 (3.1) \quad &= E(\exp^{iu[t_Y(y_0) - t_Y(y_0 - 1)]}) \dots E(\exp^{iu[t_Y(1) - t_Y(0)]}) \\
 &= E(\exp^{iuT_Y(1)})^{y_0}.
 \end{aligned}$$

Now that the problem reduces to finding the unknown variable X where $X = E(\exp^{iuT_Y(1)})$. Use $Z_y(u)$ to denote this value here, i.e., $Z_y(u) = X$. Then we can work out the expectation of $\exp^{iuT_Y(1)}$ conditional on the first y-jump time happens at t . The probability density for the jump time is then $[p_{0,-1} + p_{0,1}] \exp^{-[p_{0,-1} + p_{0,1}]t}$. If the jump happens at time t , with probability $\frac{p_{0,1}}{p_{0,-1} + p_{0,1}}$ it jumps up, in which case the expectation of $\exp^{iuT_Y(1)}$ will be $\exp^{iut} X^2$. With probability $\frac{p_{0,-1}}{p_{0,-1} + p_{0,1}}$ it jumps down and then $T_Y(1) = t$. So we derive

$$(3.2) \quad X = \int_0^\infty (p_{0,-1} + p_{0,1}) \exp^{-(p_{0,-1} + p_{0,1})t} \left(\frac{p_{0,1}}{p_{0,-1} + p_{0,1}} \exp^{iut} X^2 + \frac{p_{0,-1}}{p_{0,-1} + p_{0,1}} \exp^{iut} \right) dt.$$

To obtain X , the above integral can be solved explicitly, so we have

$$(3.3) \quad X = \frac{(p_{0,-1} + p_{0,1})}{(p_{0,-1} + p_{0,1}) - iu} \frac{p_{0,1}}{p_{0,-1} + p_{0,1}} X^2 + \frac{(p_{0,-1} + p_{0,1})}{(p_{0,-1} + p_{0,1}) - iu} \frac{p_{0,-1}}{p_{0,-1} + p_{0,1}}.$$

This is a quadratic function of X . To find which of the two possible roots is for X , we can use the condition when $u = 0$, then $X = 1$ to determine it. Thus we can obtain the following result:

$$(3.4) \quad Z_y(u) = X = \frac{1 - \sqrt{1 - 4 \frac{(p_{0,-1} + p_{0,1})}{(p_{0,-1} + p_{0,1}) - iu} \frac{p_{0,1}}{p_{0,-1} + p_{0,1}} \frac{(p_{0,-1} + p_{0,1})}{(p_{0,-1} + p_{0,1}) - iu} \frac{p_{0,-1}}{p_{0,-1} + p_{0,1}}}}{2 \frac{(p_{0,-1} + p_{0,1})}{(p_{0,-1} + p_{0,1}) - iu} \frac{p_{0,1}}{p_{0,-1} + p_{0,1}}}.$$

Similarly, we can obtain the value of $E(\exp^{iuT_X(1)})$, denoted as $Z_x(u)$. By replacing $p_{0,-1}$ and $p_{0,1}$ with $p_{-1,0}$ and $p_{1,0}$ in the right side of (3.4), we can obtain similar results for $Z_x(u)$:

$$(3.5) \quad Z_x(u) = \frac{1 - \sqrt{1 - 4 \frac{(p_{-1,0} + p_{1,0})}{(p_{-1,0} + p_{1,0}) - iu} \frac{p_{1,0}}{p_{-1,0} + p_{1,0}} \frac{(p_{-1,0} + p_{1,0})}{(p_{-1,0} + p_{1,0}) - iu} \frac{p_{-1,0}}{p_{-1,0} + p_{1,0}}}}{2 \frac{(p_{-1,0} + p_{1,0})}{(p_{-1,0} + p_{1,0}) - iu} \frac{p_{1,0}}{p_{-1,0} + p_{1,0}}}.$$

Substituting the above result into (3.1), we have the following:

$$(3.6) \quad E(\exp^{iuT_Y(y_0)}) = (Z_y(u))^{y_0}.$$

And similarly we can have

$$(3.7) \quad E(\exp^{iuT_X(x_0)}) = (Z_x(u))^{x_0}.$$

The characteristic function of $T_Y(y_0) - T_X(x_0)$ can be obtained so far:

$$(3.8) \quad E(\exp^{iu[T_Y(y_0) - T_X(x_0)]}) = (Z_y(u))^{y_0} (Z_x(-u))^{x_0}.$$

Thus the probability distribution of $T_Y(y_0) - T_X(x_0)$ can be derived:

$$(3.9) \quad \begin{aligned} p(T_Y(y_0) - T_X(x_0)) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} E(\exp^{iu[T_Y(y_0) - T_X(x_0)]}) \exp^{-iut} du \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} (Z_y(u))^{y_0} (Z_x(-u))^{x_0} \exp^{-iut} du. \end{aligned}$$

Also by the inversion theorem of Gil-Pelaez [5], we can obtain the cumulative function of $T_Y(y_0) - T_X(x_0)$ from its characteristic function. Here we use a particular form of this inversion theorem, which is given by equation (3.9) of [12]:

$$(3.10) \quad p(T_Y(y_0) - T_X(x_0) \leq x) = \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp^{-iux} [(Z_y(u))^{y_0} (Z_x(-u))^{x_0}]}{iu} du.$$

Recall that what we are interested in is the probability of the random walk hitting the X-axis before the Y-axis, which equals the probability of $p(T_Y(y_0) - T_X(x_0) \leq 0)$. Thus by evaluating the equation above at $x = 0$, we get the following result:

$$(3.11) \quad \begin{aligned} p(T_Y(y_0) - T_X(x_0) \leq 0) &= \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(Z_y(u))^{y_0} (Z_x(-u))^{x_0}}{iu} du \\ &= \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} \frac{\Im(Z_y(u))^{y_0} (Z_x(-u))^{x_0}}{u} du. \end{aligned}$$

with $Z_y(u)$ and $Z_x(u)$ defined in (3.4) and (3.5).

Thus we have derived the probability of mid-price increase starting at x_0 and y_0 orders at the bid and ask sides, respectively. Let the price-increase probability be $p_{inc}(x_0, y_0)$, then

$$(3.12) \quad p_{inc}(x_0, y_0) = p(T_Y(y_0) - T_X(x_0) \leq 0) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} \frac{\Im(Z_y(u))^{y_0} (Z_x(-u))^{x_0}}{u} du,$$

with $Z_y(u)$ and $Z_x(u)$ defined in (3.4) and (3.5). This is our final result for the price increase probability $p_{inc}(x_0, y_0)$.

The integration part of (3.12) can be evaluated easily using numerical integration methods.

4. Conclusion. In this paper we showed some errors made in Proposition 3 of [2] in trying to obtain an explicit solution of the probability of a mid-price increase given the state of the order book under an unbalanced order flow. Some of the errors can be easily corrected. However, a serious error occurs in the proof, on which the whole of Proposition 3 falls, and which occurs due to the citing of incorrect results from another paper [13].

To solve the problem, we demonstrated an explicit solution using characteristic functions of these conditional probability distributions. The solution obtained in our approach has a simple form and can be evaluated easily and fast. Therefore, this solution can be used to calculate the probabilities of the dynamics of price increase in the model described in [2]. This formula can be used as the foundation of empirical studies of the LOB as well as further theoretic research on price dynamics of the LOB driven by OFI under certain assumptions.

REFERENCES

- [1] MARCO AVELLANEDA, JOSH REED, AND SASHA STOIKOV, *Forecasting prices from level- i quotes in the presence of hidden liquidity*, *Algorithmic Finance*, 1 (2011), pp. 35–43.
- [2] RAMA CONT AND ADRIEN DE LARRARD, *Price dynamics in a Markovian limit order market*, *SIAM Journal on Financial Mathematics*, 4 (2013), pp. 1–25.
- [3] RAMA CONT, ARSENIY KUKANOV, AND SASHA STOIKOV, *The price impact of order book events*, *Journal of financial econometrics*, 12 (2013), pp. 47–88.
- [4] RAMA CONT, SASHA STOIKOV, AND RISHI TALREJA, *A stochastic model for order book dynamics*, *Operations research*, 58 (2010), pp. 549–563.
- [5] J GIL-PELAEZ, *Note on the inversion theorem*, *Biometrika*, 38 (1951), pp. 481–482.
- [6] MARTIN D. GOULD, MASON A. PORTER, STACY WILLIAMS, MARK McDONALD, DANIEL J. FENN, AND SAM D. HOWISON, *Limit order books*, *Quantitative Finance*, 13 (2013), pp. 1709–1742.
- [7] DONALD GROSS AND CARL HARRIS, *Fundamentals of queueing theory*, (1998).
- [8] OLIVIER GUÉANT, CHARLES-ALBERT LEHALLE, AND JOAQUIN FERNANDEZ-TAPIA, *Optimal portfolio liquidation with limit orders*, *SIAM Journal on Financial Mathematics*, 3 (2012), pp. 740–764.

- [9] FABIEN GUILBAUD AND HUYEN PHAM, *Optimal high-frequency trading with limit and market orders*, Quantitative Finance, 13 (2013), pp. 79–94.
- [10] V YU KOROLEV, AV CHERTOK, A YU KORCHAGIN, AND AI ZEIFMAN, *Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes*, Applied Mathematics and Computation, 253 (2015), pp. 224–241.
- [11] I KURKOVA AND KILIAN RASCHEL, *Random walks in $(\mathbb{Z}_+)^2$ with non-zero drift absorbed at the axes*, Bulletin de la Société Mathématique de France, 139 (2011), pp. 341–388.
- [12] MARTIN SCHMELZLE, *Option pricing formulae using Fourier transform: Theory and application*, Preprint, <http://pfadintegral.com>, (2010).
- [13] JOHAN SH VAN LEEUWAARDEN, KILIAN RASCHEL, ET AL., *Random walks reaching against all odds the other side of the quarter plane*, Journal of Applied Probability, 50 (2013), pp. 85–102.